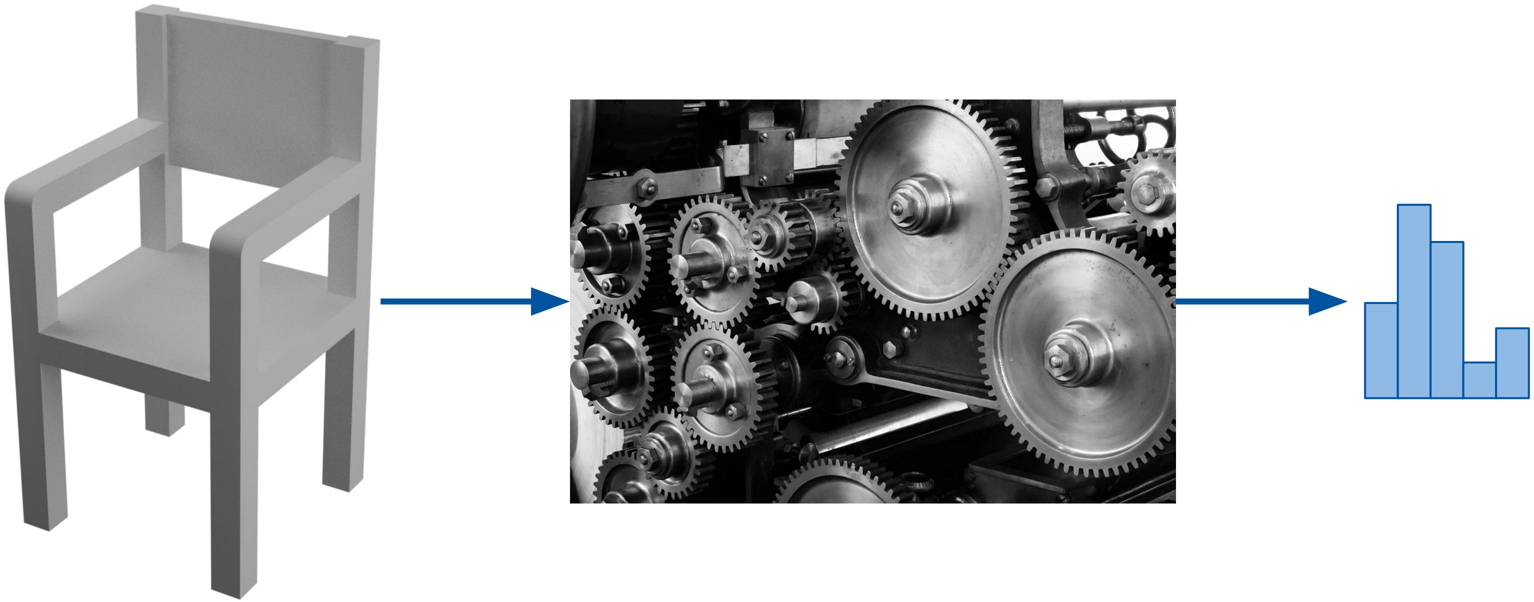
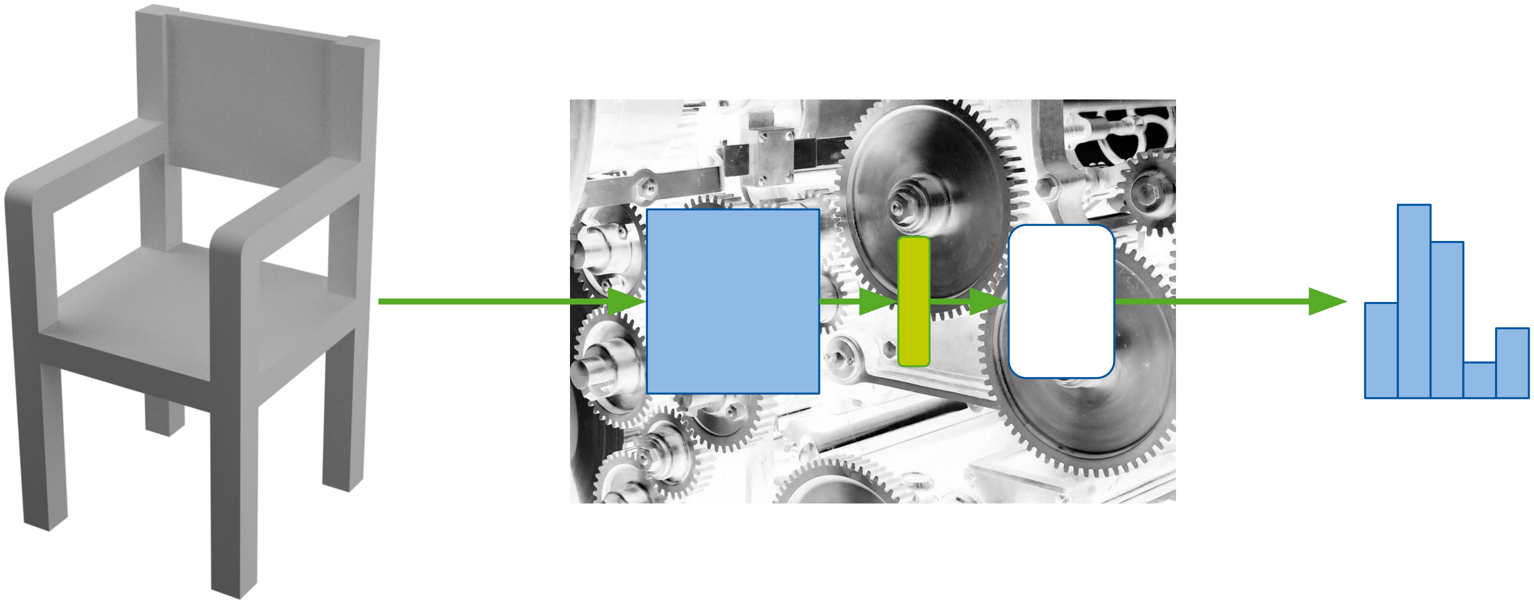
Learned Embeddings for Geometric Data

Visual Computing Institute
RWTH AACHEN UNIVERSITY

Encodings of Geometric Data



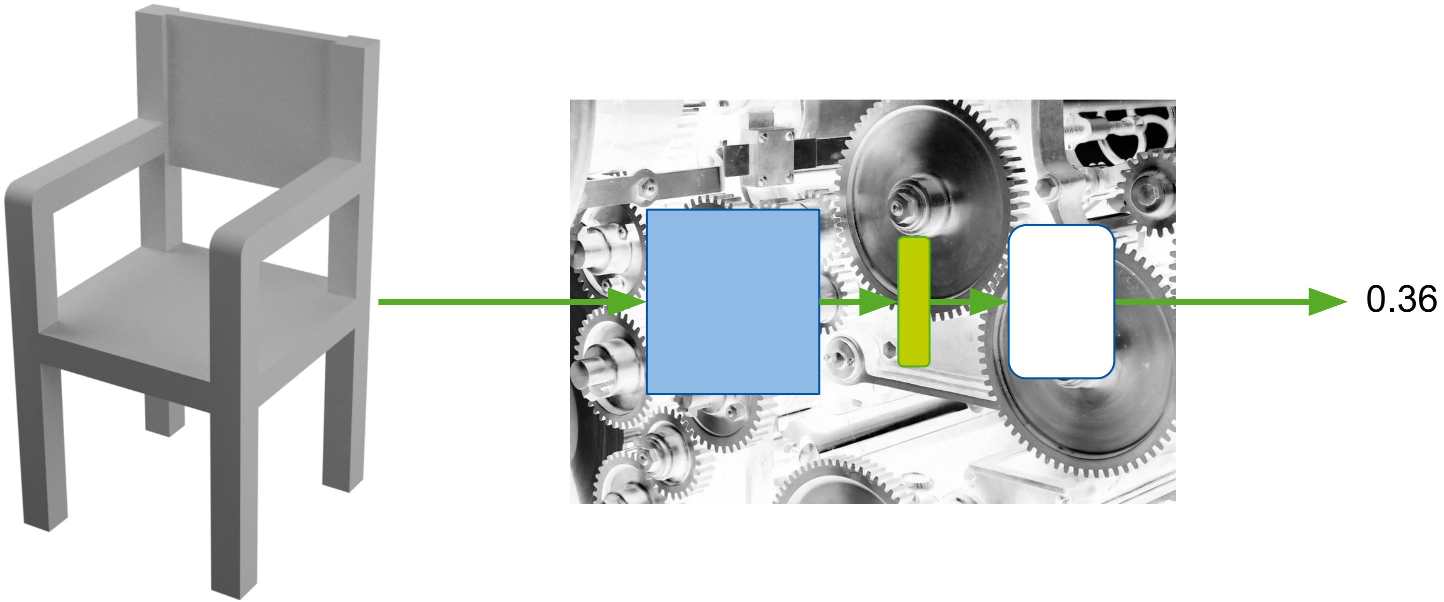
Encodings of Geometric Data



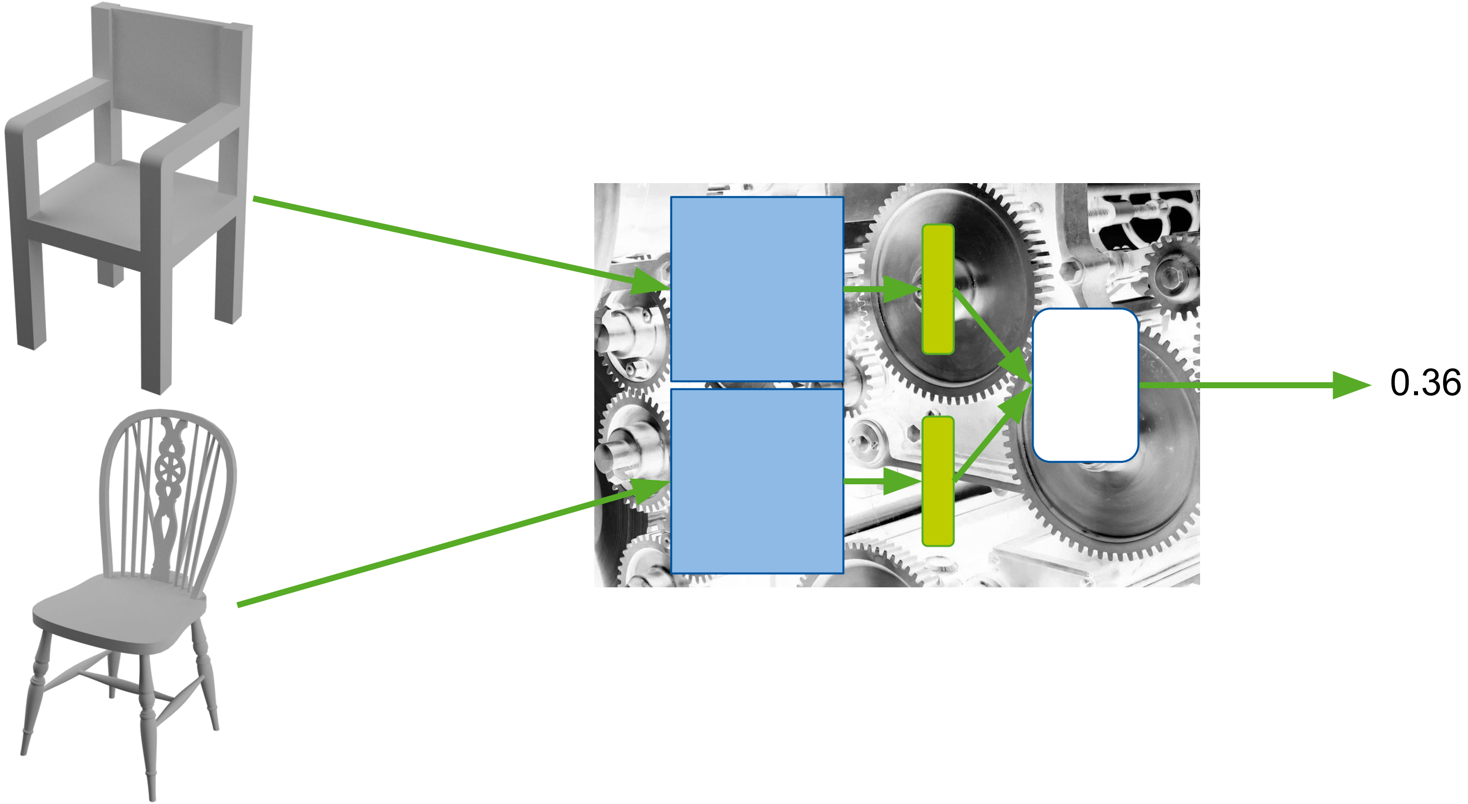
Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

Visual Computing Institute
RWTH AACHEN UNIVERSITY

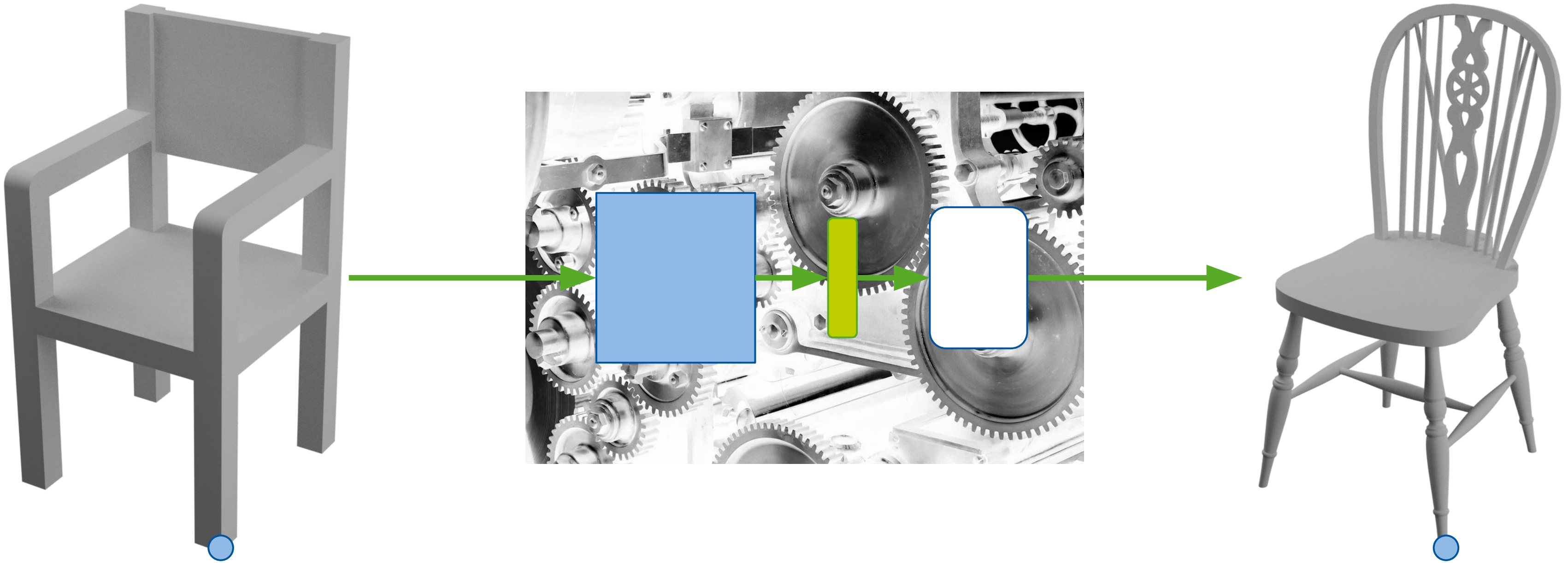
Encodings of Geometric Data



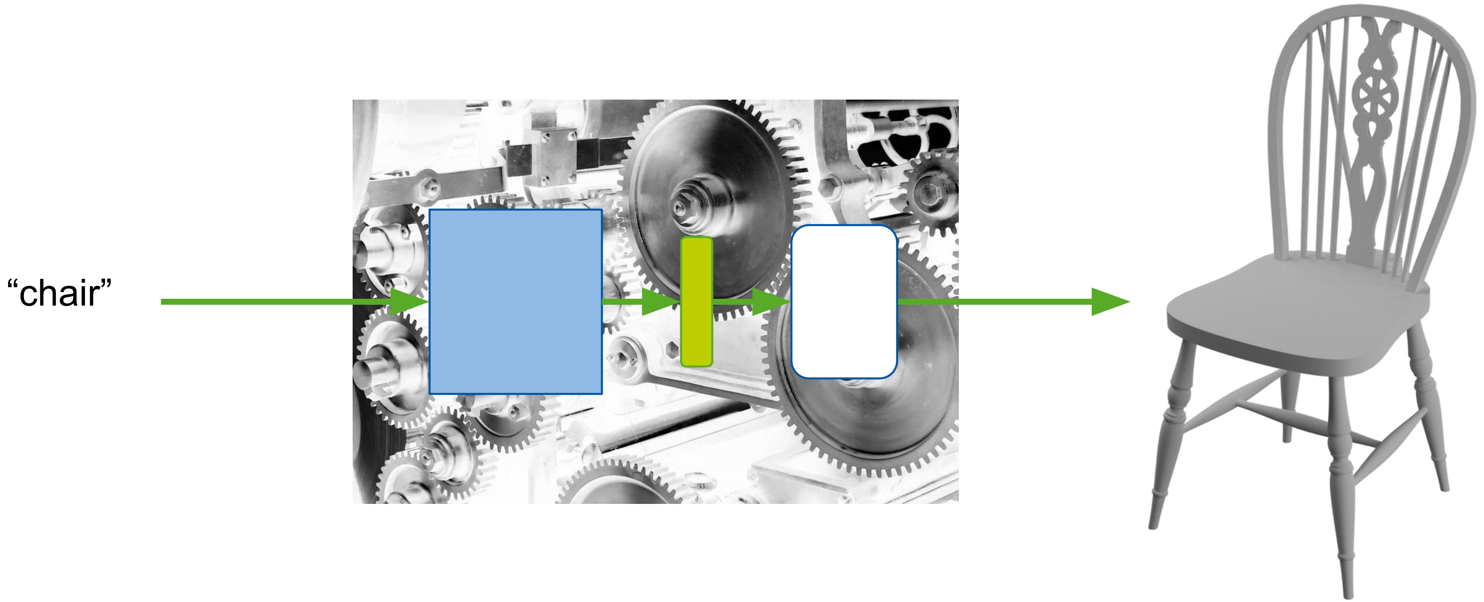
Encodings of Geometric Data



Encodings of Geometric Data



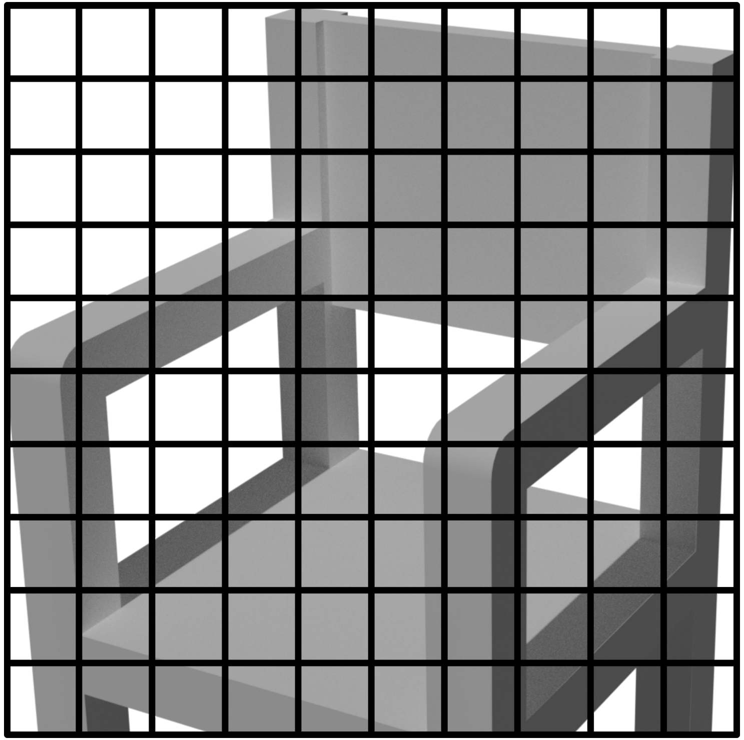
Encodings of Geometric Data

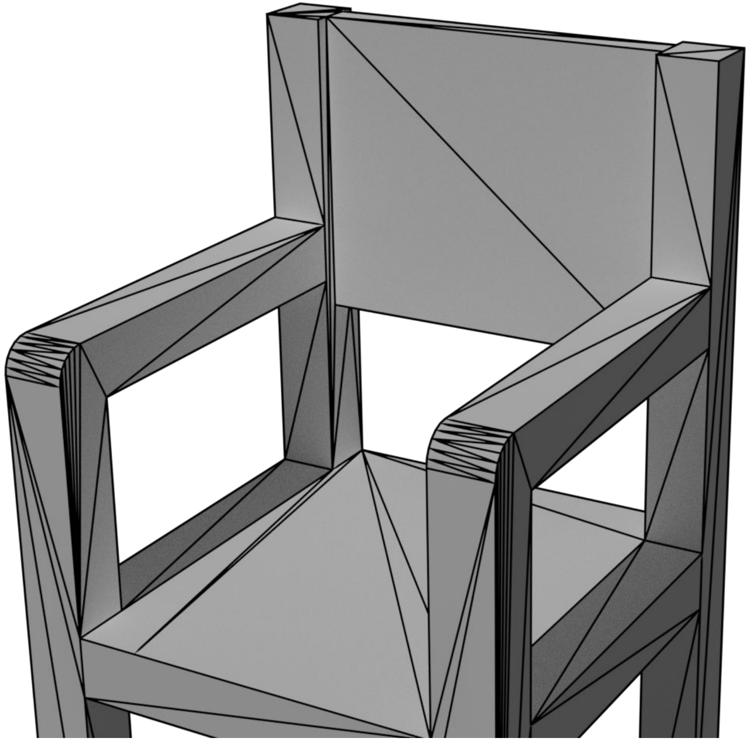


Key Questions

- How can we compute encodings of geometric data?
- How can we design the embedding space?

Representations of Geometric Data





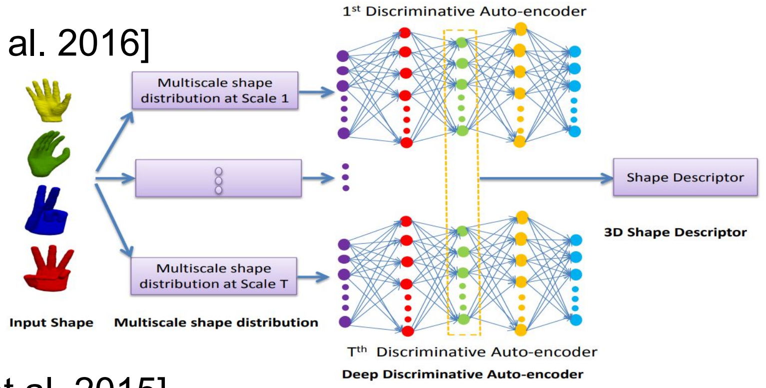
Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surface Patches

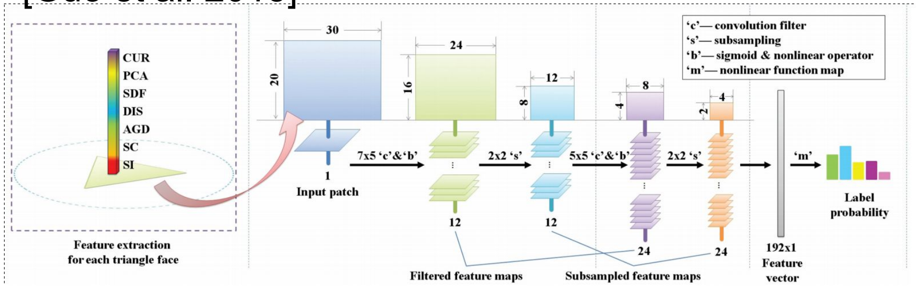
Input Representations for Geometric Data

- **Handcrafted Descriptors**
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surface Patches

[Xie et al. 2016]



[Guo et al. 2015]



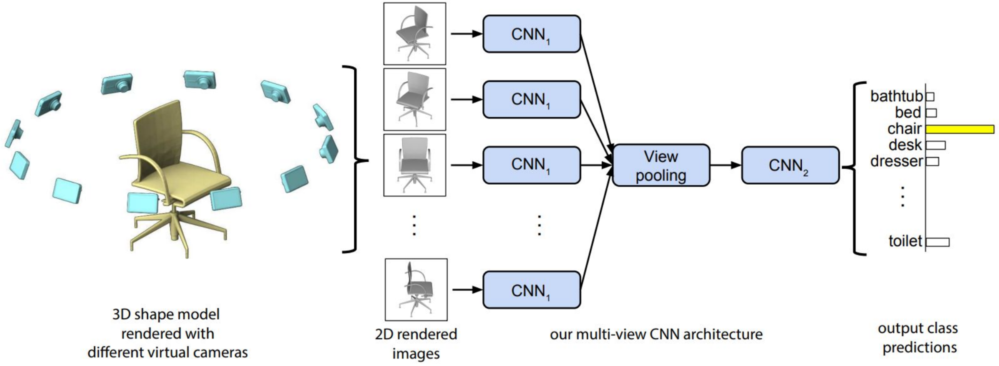
Learning on Handcrafted Descriptors

- Introduce domain knowledge by constructing descriptors
  - Less data
- Limited by the expressiveness of descriptors

Input Representations for Geometric Data

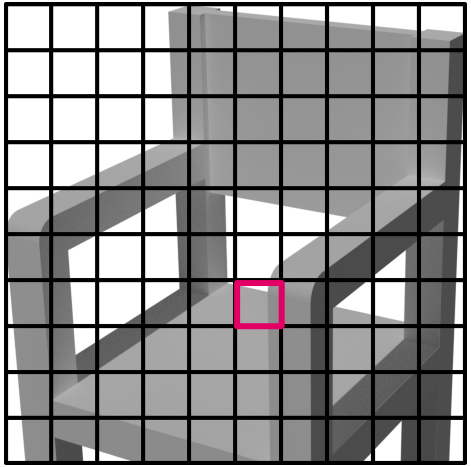
- Handcrafted Descriptors
- **Images**
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surface Patches

[Su et al. 2015]



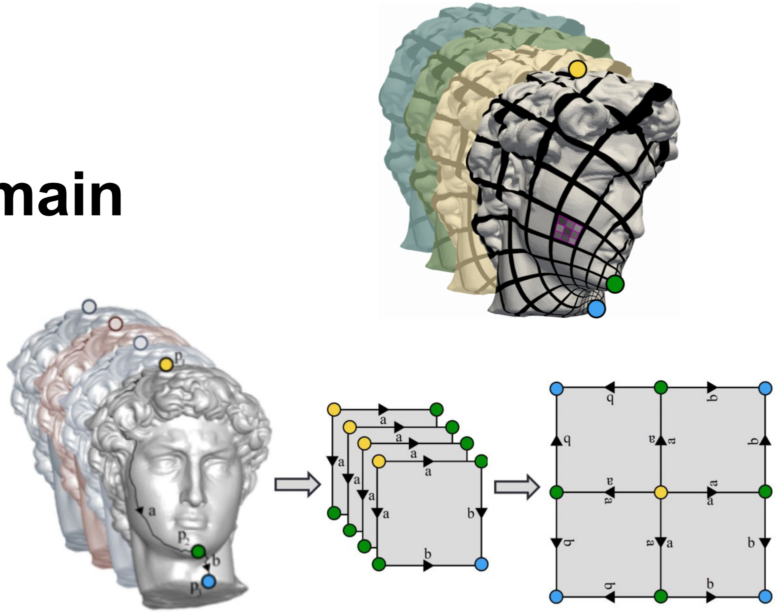
Learning on Images

- Available prior work on 2D CNNs
  - Pretrained weights
  - Availability of datasets
- View selection is challenging
- Convolution operation does not consider geodesic neighbourhood



Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- **2D Parametrization Domain**
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surface Patches



[Maron et al. 2017]

Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

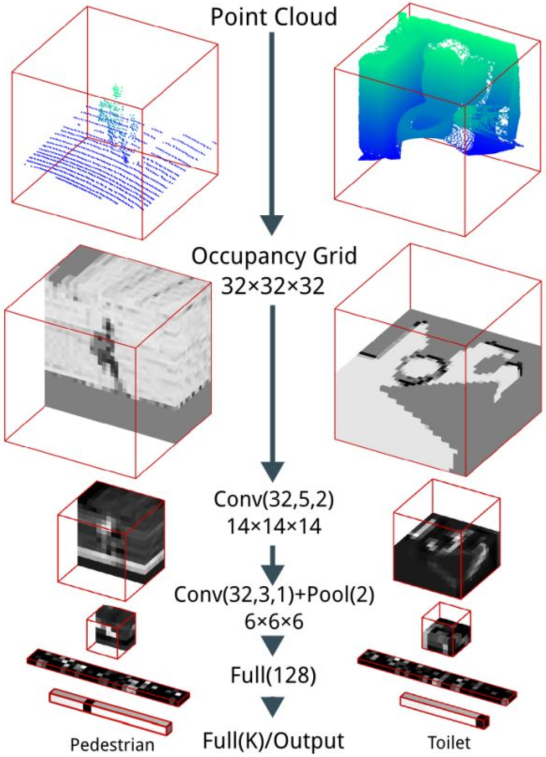
Learning in the Parametrization Domain

- 2D CNNs can be used on the surface
- Limitations in the quality of available (genus 0) datasets
- Some sensitivity to the chosen cut graph vertices

Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- **Voxel Grids**
- Point Clouds
- Graphs
- Curved Surface Patches

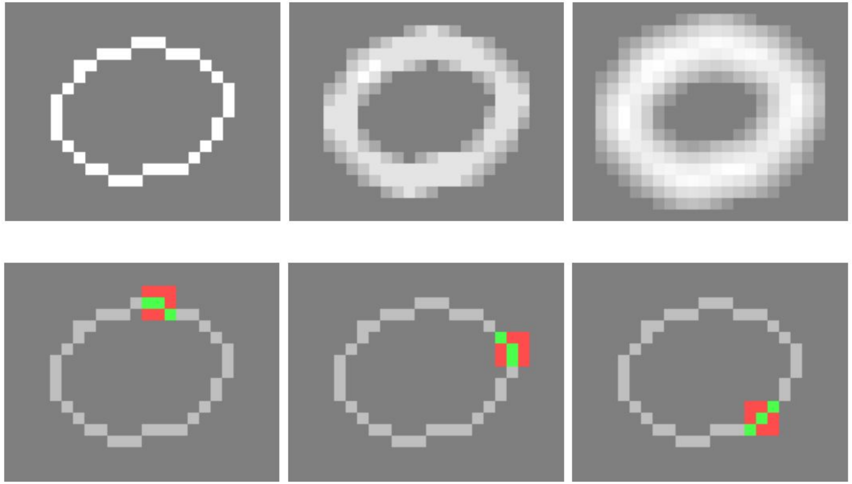
[Maturana et al. 2015]



Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- **Voxel Grids**
- Point Clouds
- Graphs
- Curved Surface Patches

[Graham et al. 2018]

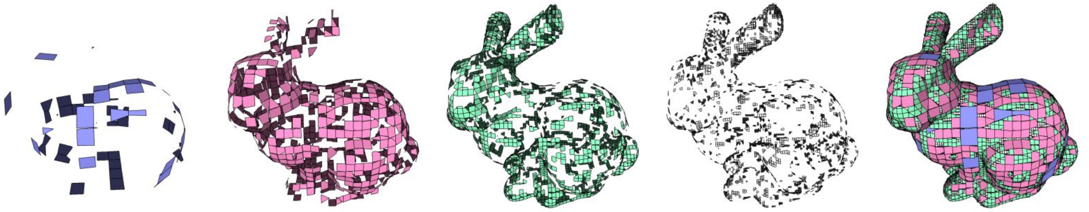


Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- **Voxel Grids**
- Point Clouds
- Graphs
- Curved Surface Patches

[Wang et al. 2018]



(a) 4th-level (b) 5th-level (c) 6th-level (d) 7th-level (e) all non-empty leaf nodes

Learning on Voxel Grids

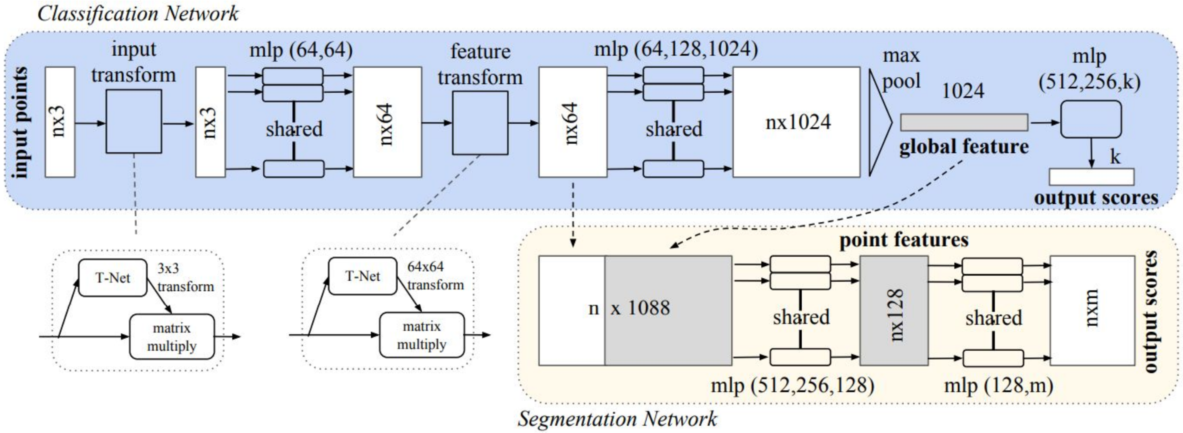
- Straight-forward extension of 2D CNNs to 3D CNNs possible
  - Grid structure
  - Euclidean domain
- Aliasing artefacts due to grid alignment
- Memory efficiency requires custom operations

Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- **Point Clouds**
- Graphs
- Curved Surface Patches

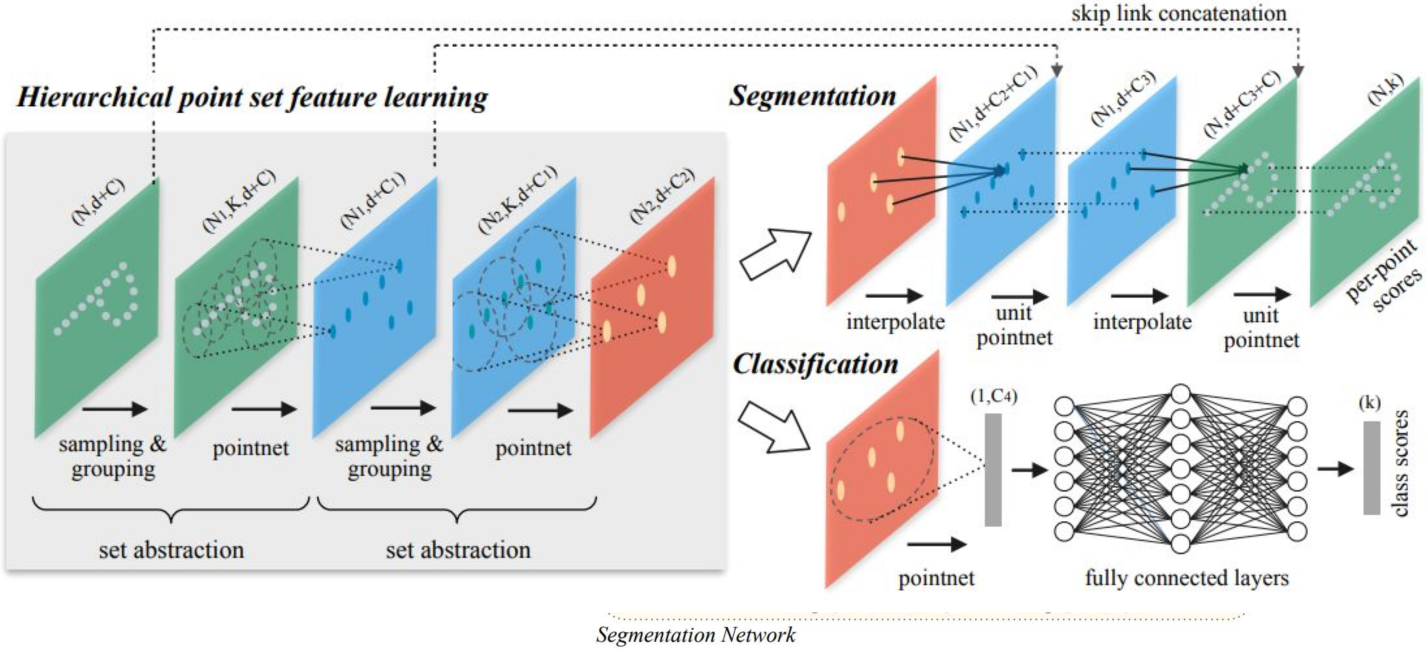
Deep Sets & PointNet

- ensure permutation invariance by applying the same function (MLP) to each element of a set
- aggregate information by applying a symmetric function (e.g. maximum) to all transformed elements of the set



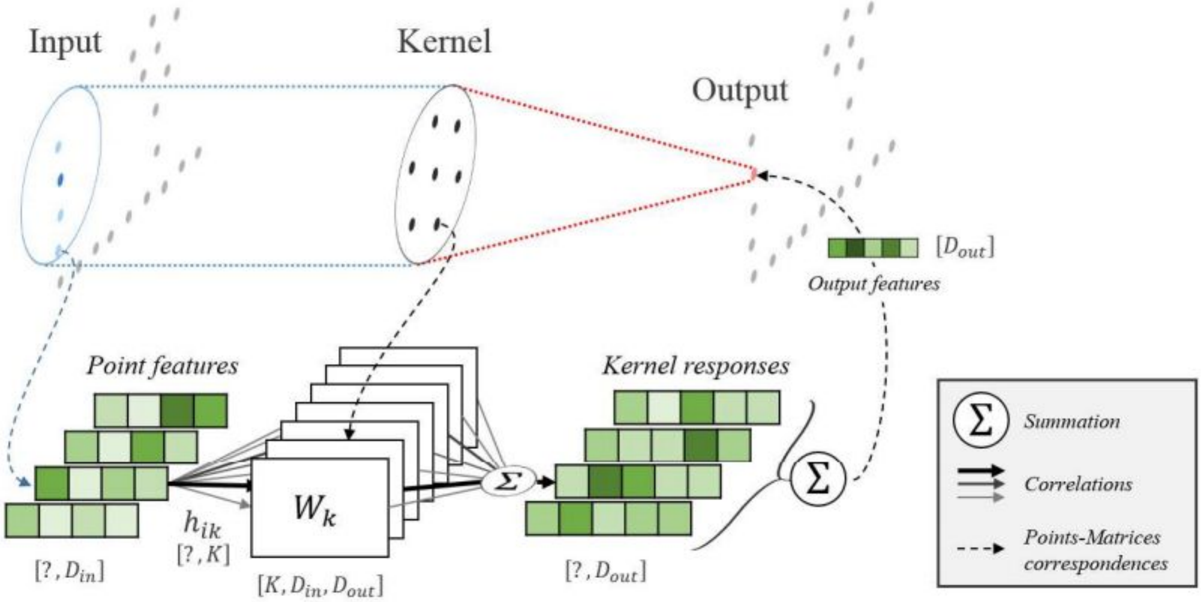
PointNet++

- Learn features at various scales



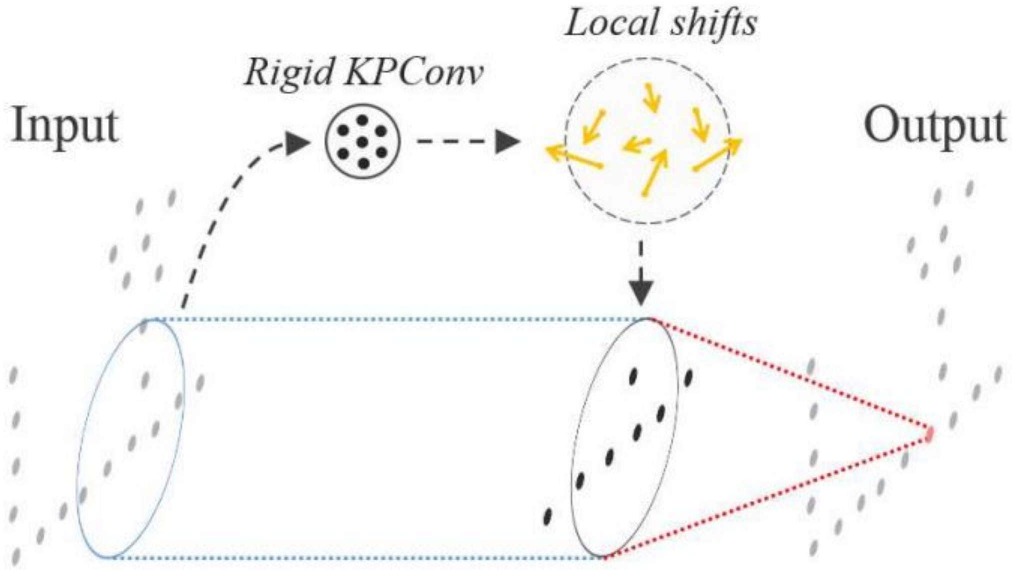
Kpconv [Thomas et al. 2019]

- Interpolate local point features at fixed number of kernel points



Kpconv [Thomas et al. 2019]

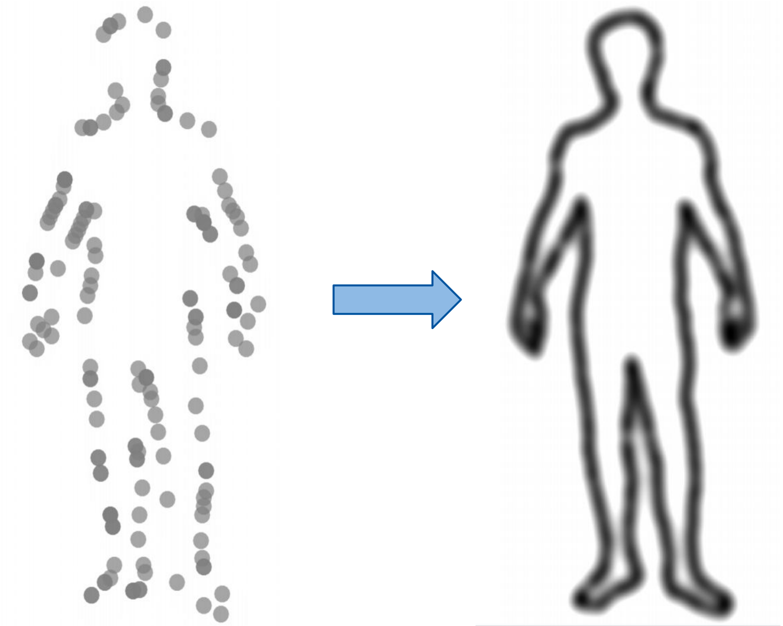
- Kernel point positions can also be learned via local shift vectors



Point convolutional neural networks by extension operators

- Extend local point features (f) to 3D volume with Gaussians
  - weighted (w) inversely by local density

[Atzmon et al. 2018]



Learning on Point Clouds

- Robust w.r.t. the quality of the dataset
- Ignores existing connectivity information
- Inefficient representation
  - Planar regions have to be sampled

Input Representations for Geometric Data

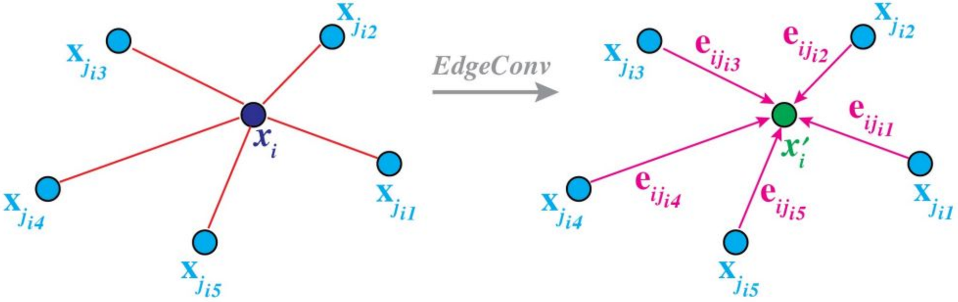
- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- **Graphs**
- Curved Surface Patches

[Wang et al. 2018]

Input Representations for Geometric Data

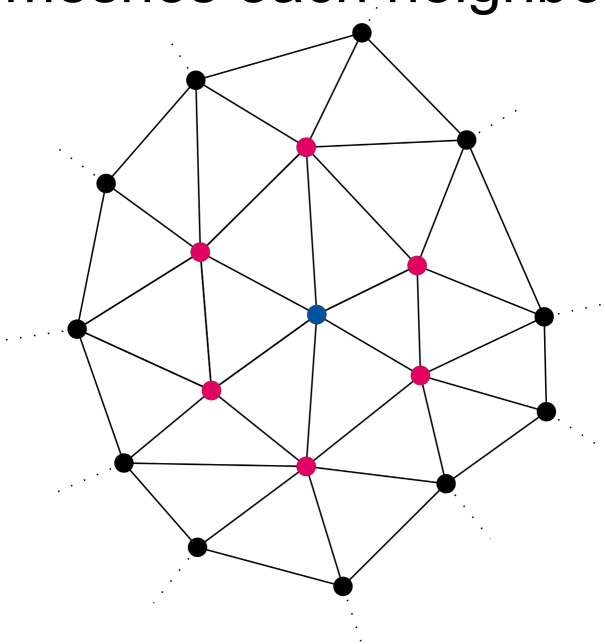
- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- **Graphs**
- Curved Surface Patches

[Wang et al. 2019]



Learning on Graphs for Geometric Data

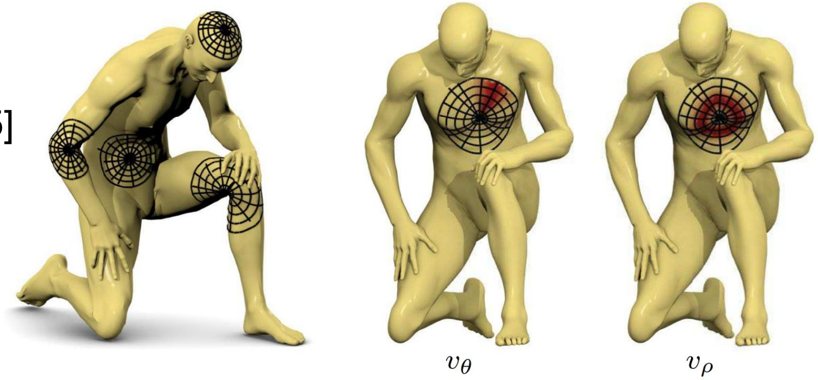
- Flexibel in terms of the input quality
  - Point clouds
  - (non-manifold) meshes
- Even for manifold meshes each neighbour is treated the same



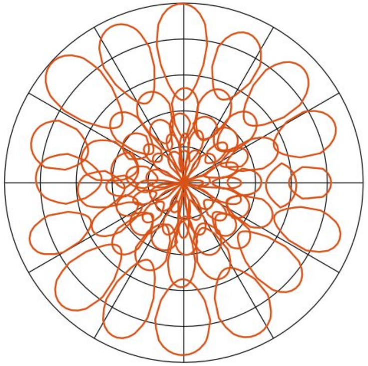
Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- **Curved Surface Patches**

[Masci et al. 2015]



[Monti et al. 2017]



Learning on Curved Surface Patches

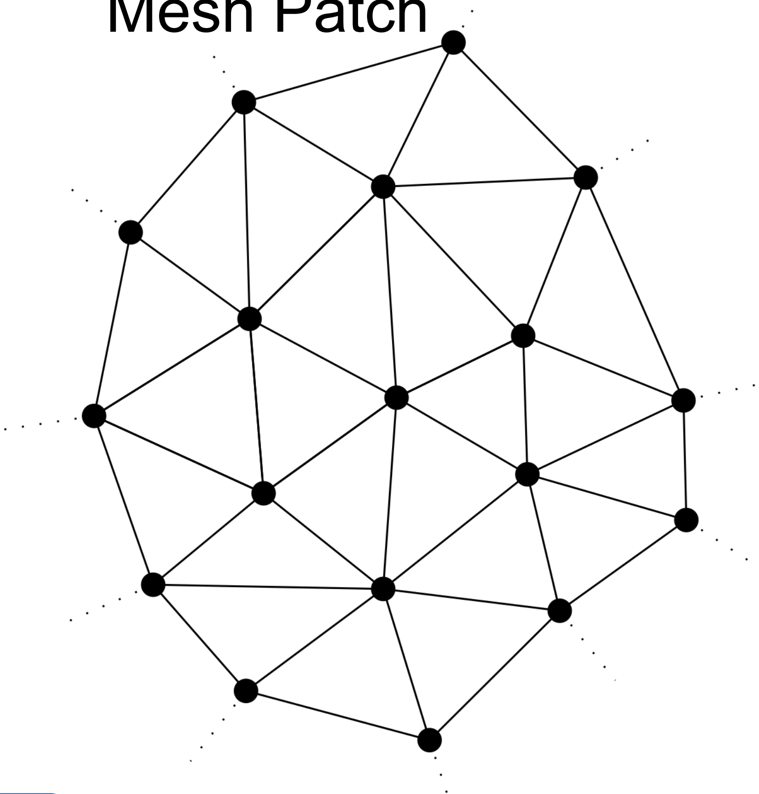
- Convolution operation on curved surfaces that respects geodesic neighbourhood
- (some) input requirements
- Orientation of filters is not straight-forward

Our contribution

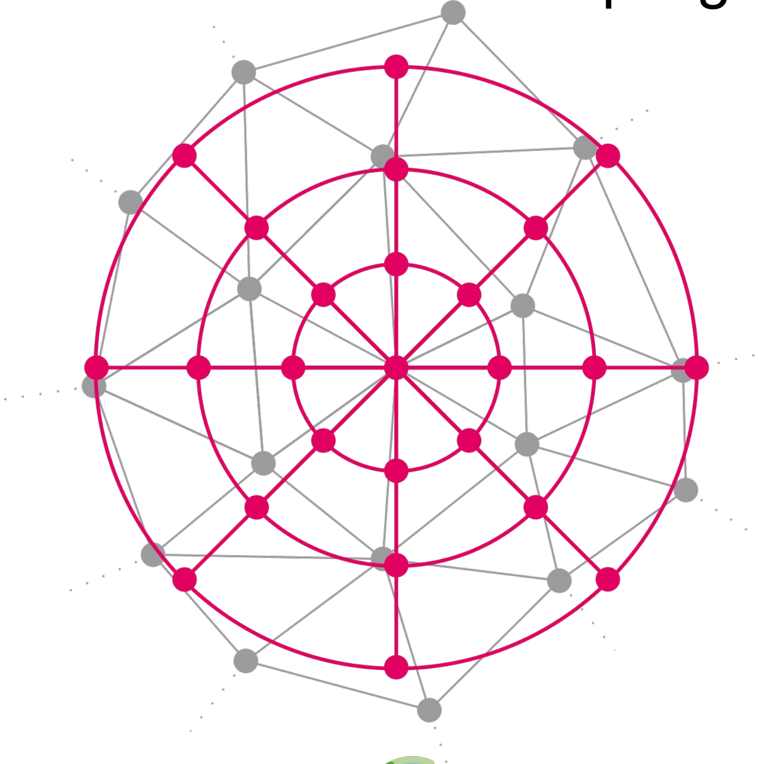
A Simple Approach to Intrinsic Correspondence Learning on Unstructured 3D Meshes

Learning on Meshes: SpiralNet (and subsequent works)

Mesh Patch



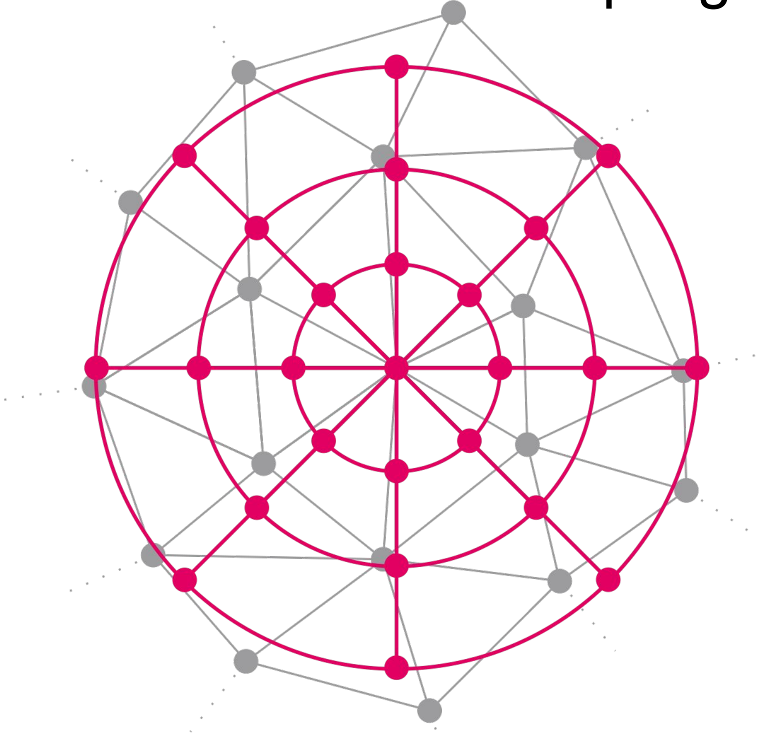
Kernel based resampling



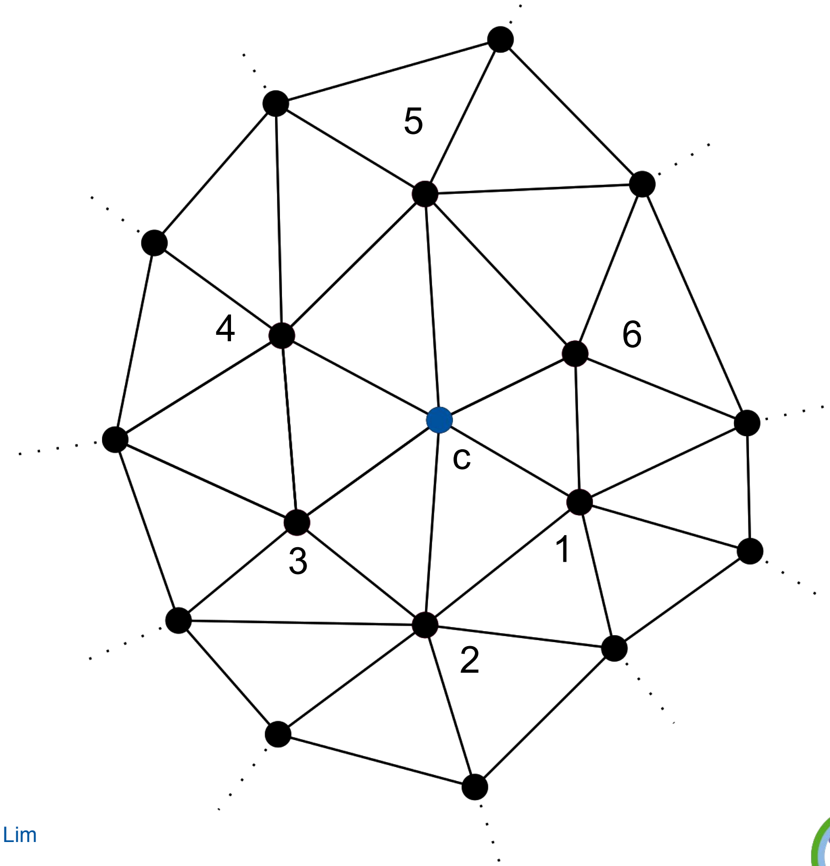
Learning on Meshes: SpiralNet (and subsequent works)

- Loss of data fidelity
- Unclear interpolation operation
- Pre-processing costs

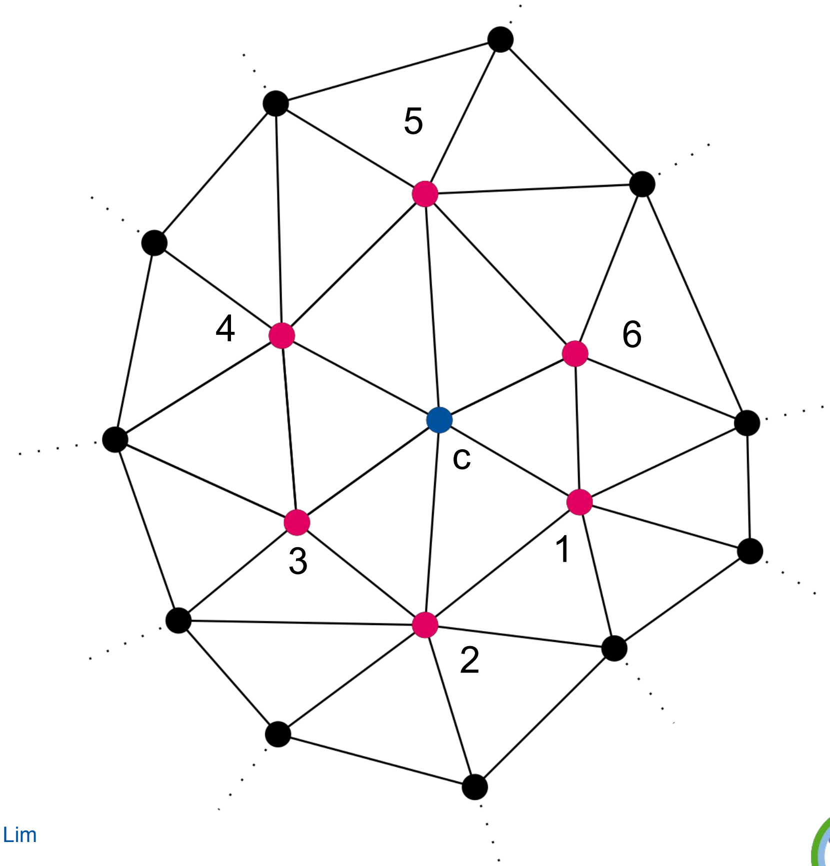
Kernel based resampling



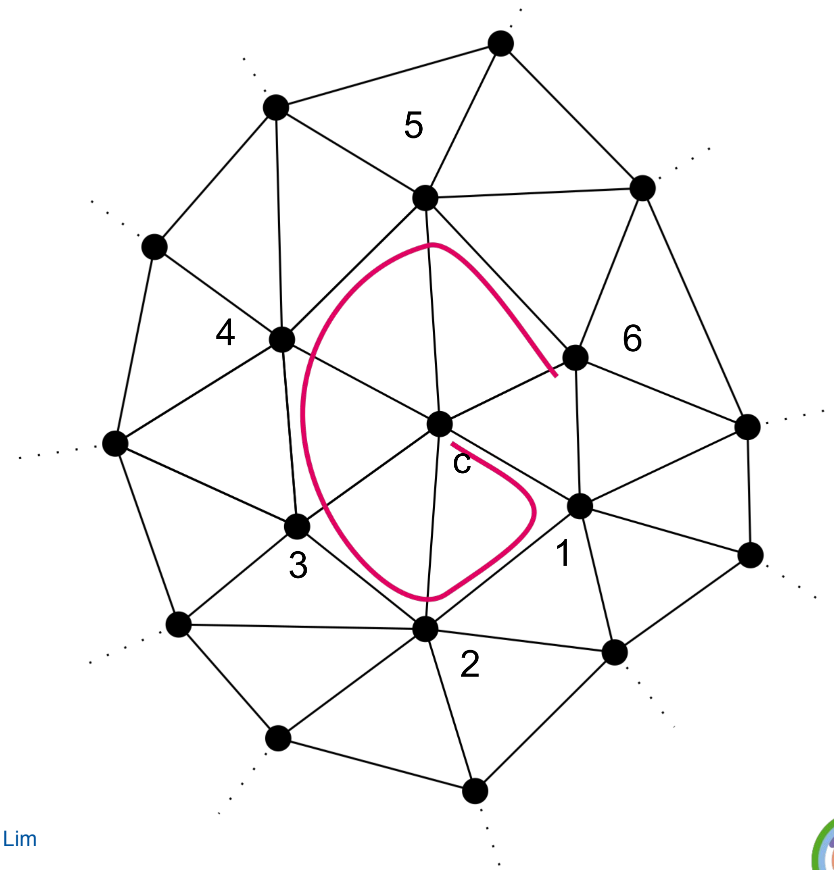
Spiral Operator



Spiral Operator

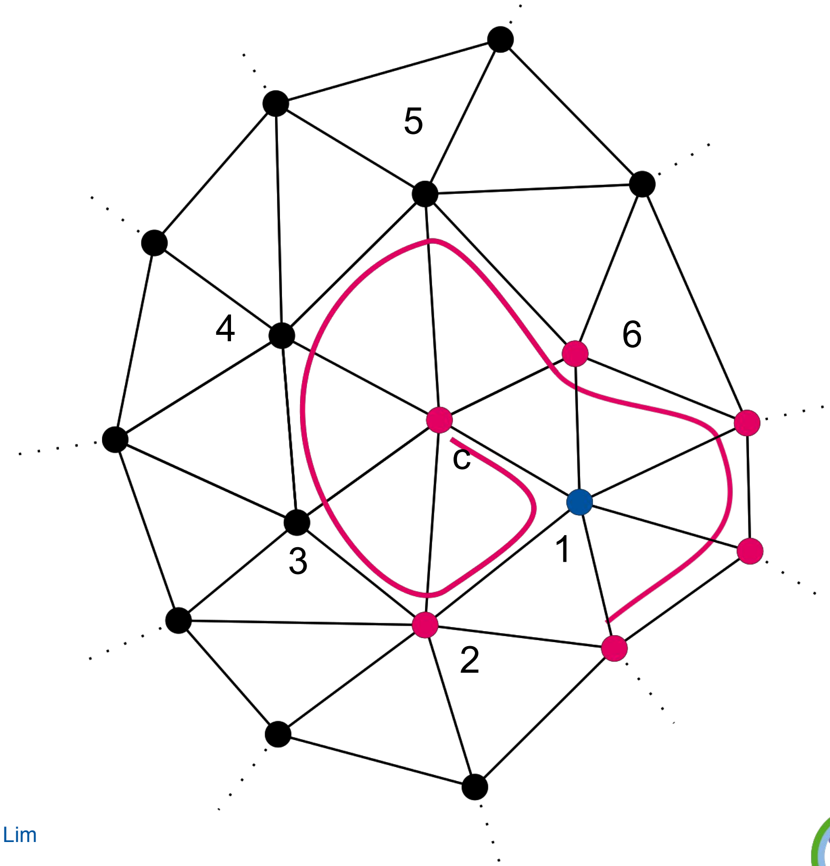


Spiral Operator

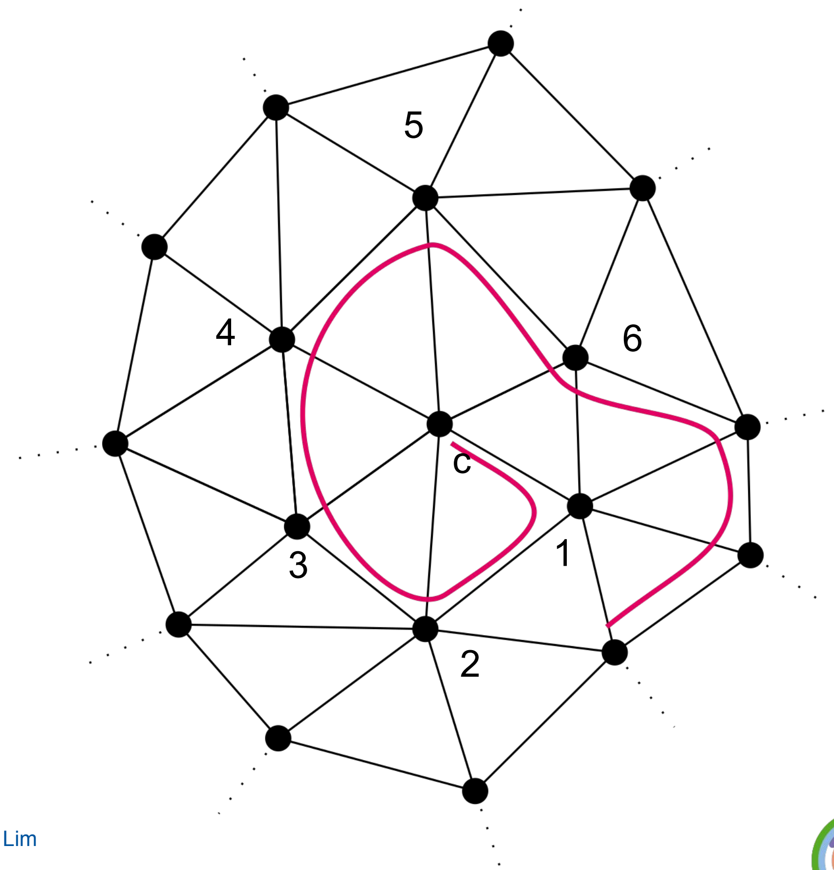


Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

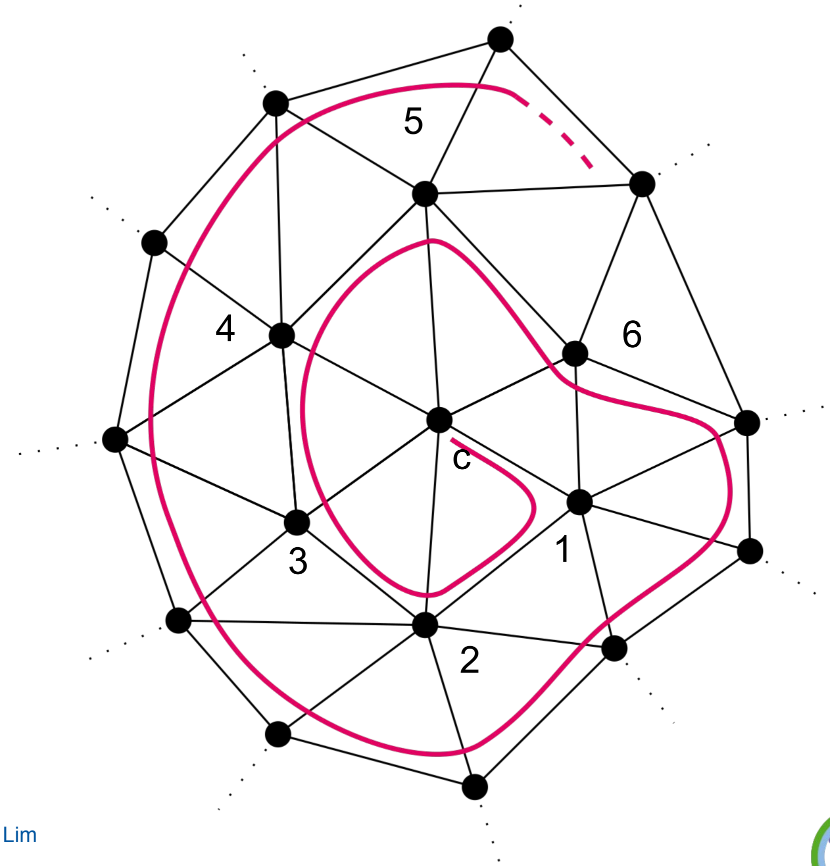
Spiral Operator



Spiral Operator



Spiral Operator



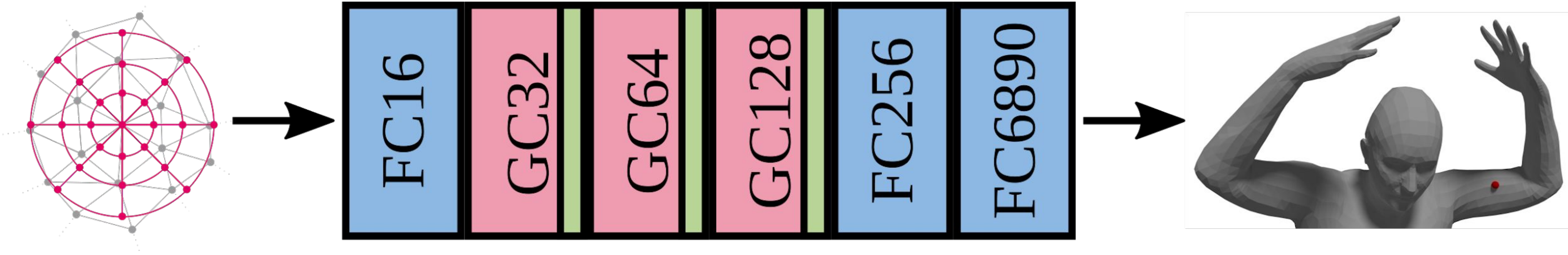
Neighbourhood Encoding

- seq = [c, 1, 2, 3, ...]
- f(seq) = [f(c), f(1), f(2), f(3), ...]

Possible encodings:

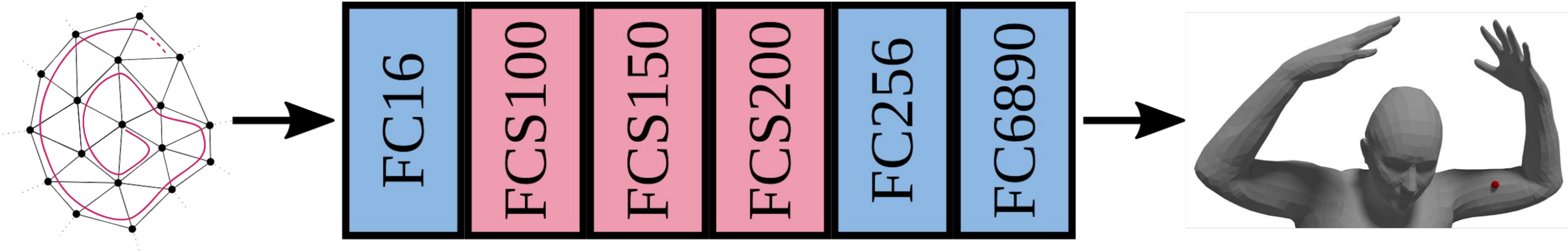
- h = LSTM(f(seq))
- h = FC(f(seq))
- h = g * f(seq) [Bokhnyak et al.]

GCNN

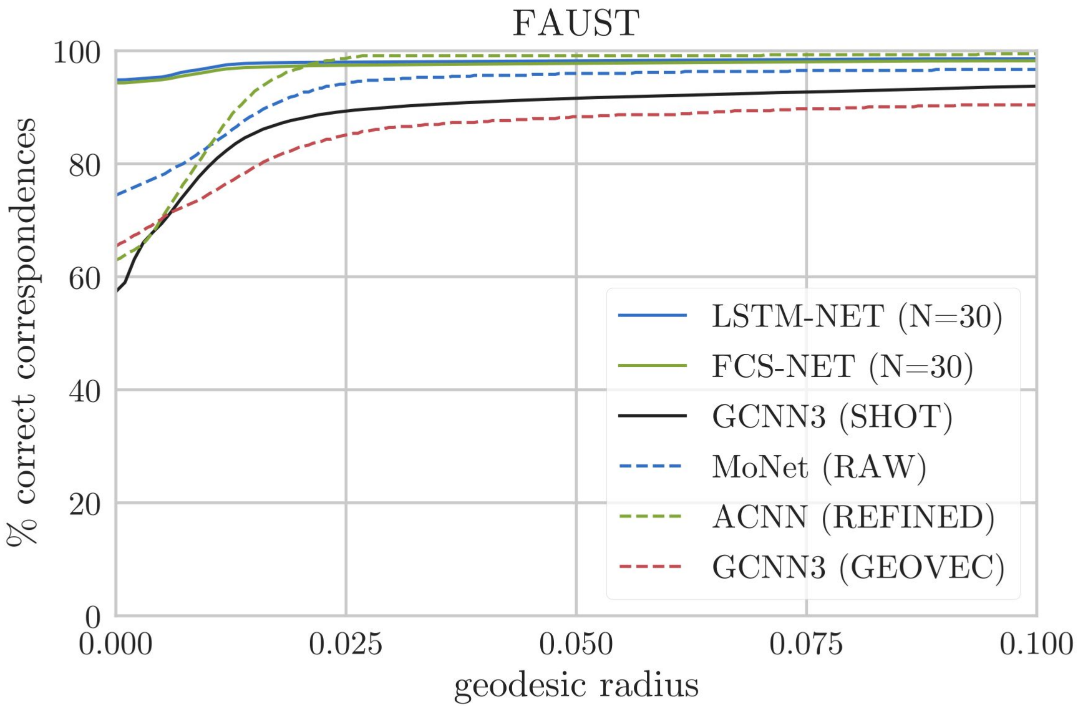


LSTM-Net

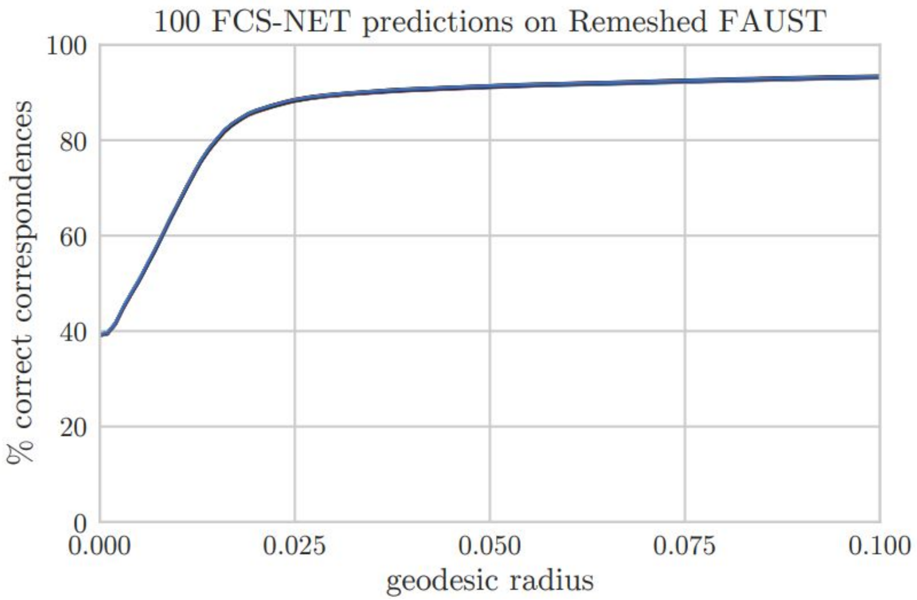
FCS-NET

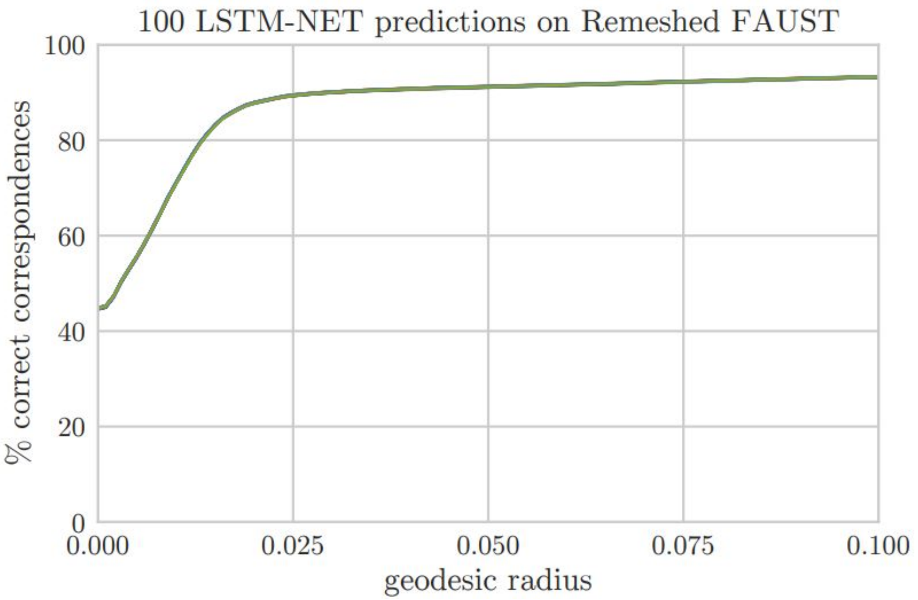


Results


FAUST
100
80
60
40
20
0
0.000
0.025
0.050
0.075
0.100
geodesic radius
% correct correspondences
LSTM-NET (N=30)
FCS-NET (N=30)
GCNN3 (SHOT)
MoNet (RAW)
ACNN (REFINED)
GCNN3 (GEOVEC)

Robustness to different Rotations





Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surfaces

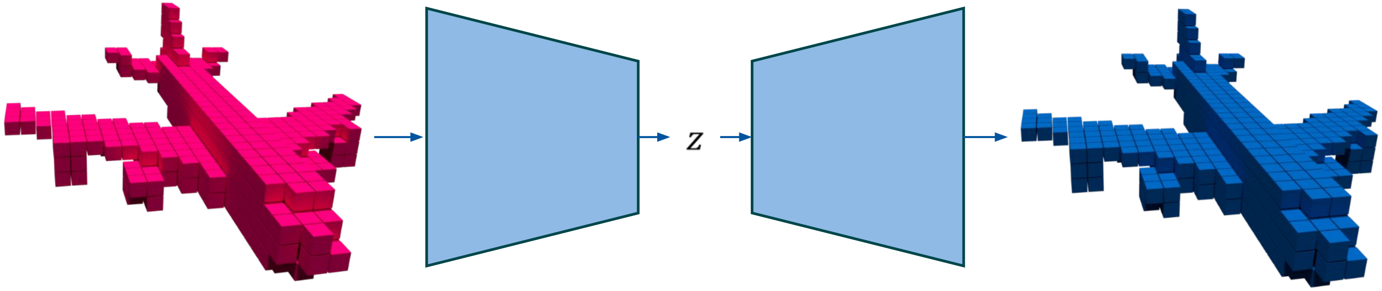
Input Representations for Geometric Data

- Handcrafted Descriptors
- Images
- 2D Parametrization Domain
- Voxel Grids
- Point Clouds
- Graphs
- Curved Surfaces
- **Combinations thereof**

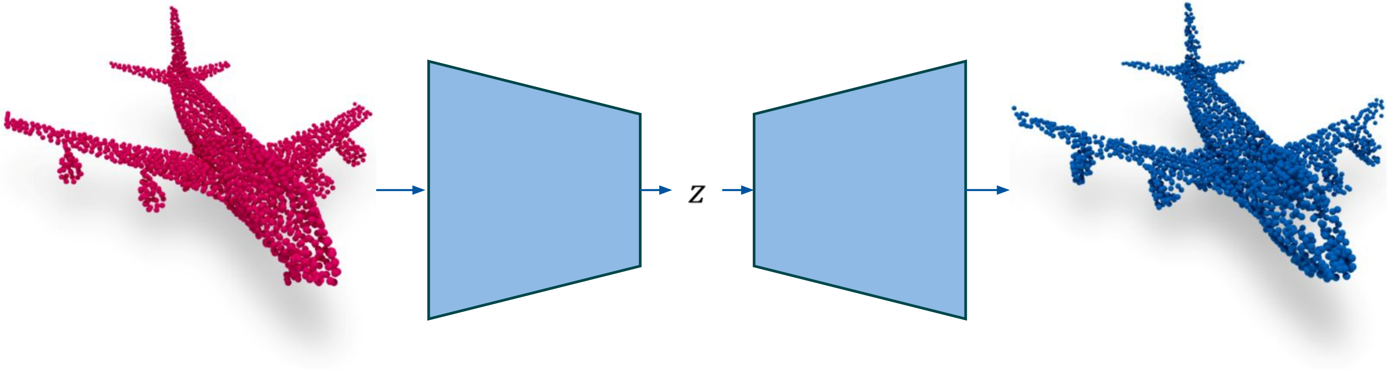
Our contribution

A Convolutional Decoder for Point Clouds using Adaptive Instance Normalization

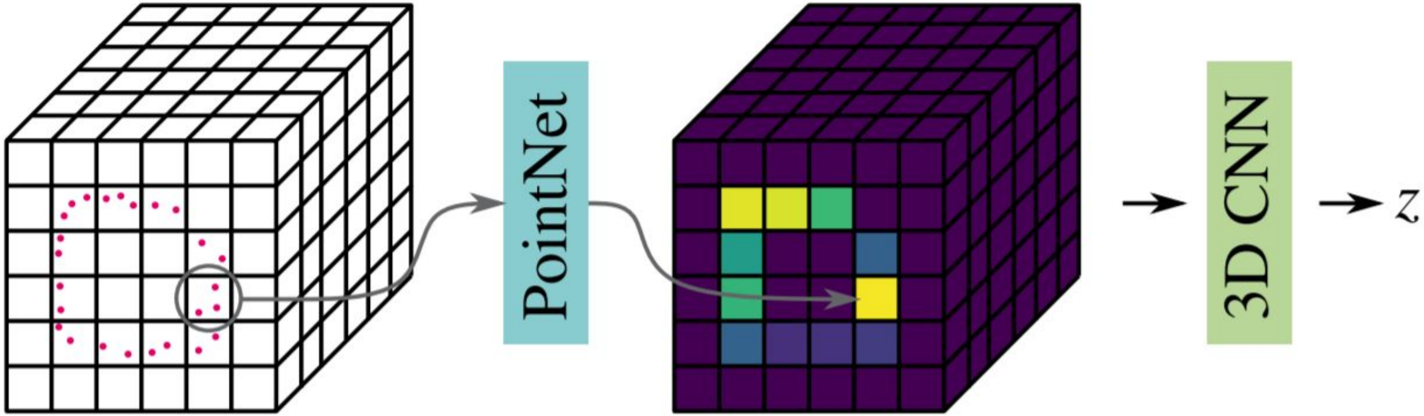
Autoencoding



Autoencoding



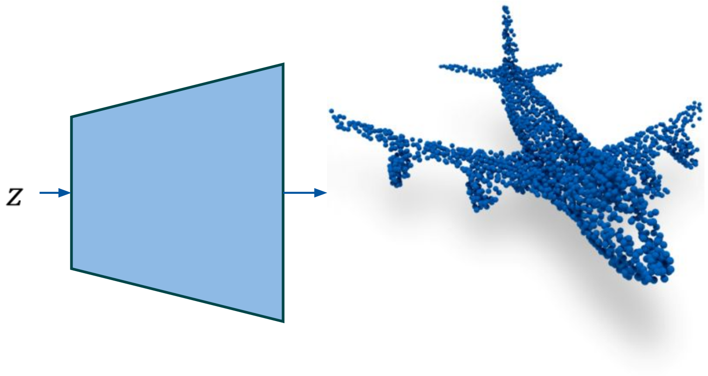
Encoder



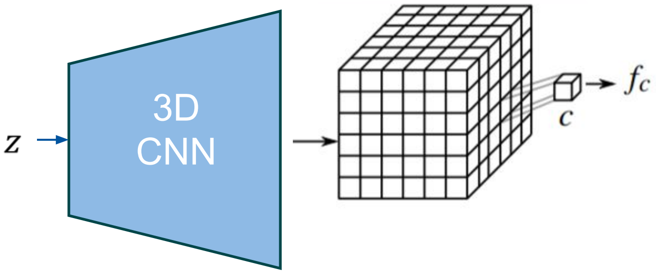
Similar to “Fully-convolutional point networks for large-scale point clouds”

Rethage et al., 2018

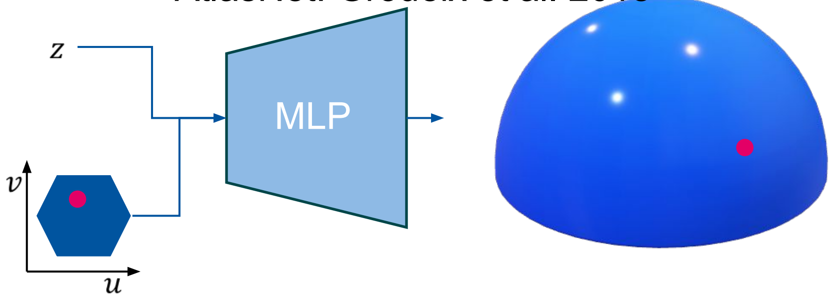
Decoding



Learning Localized Maps

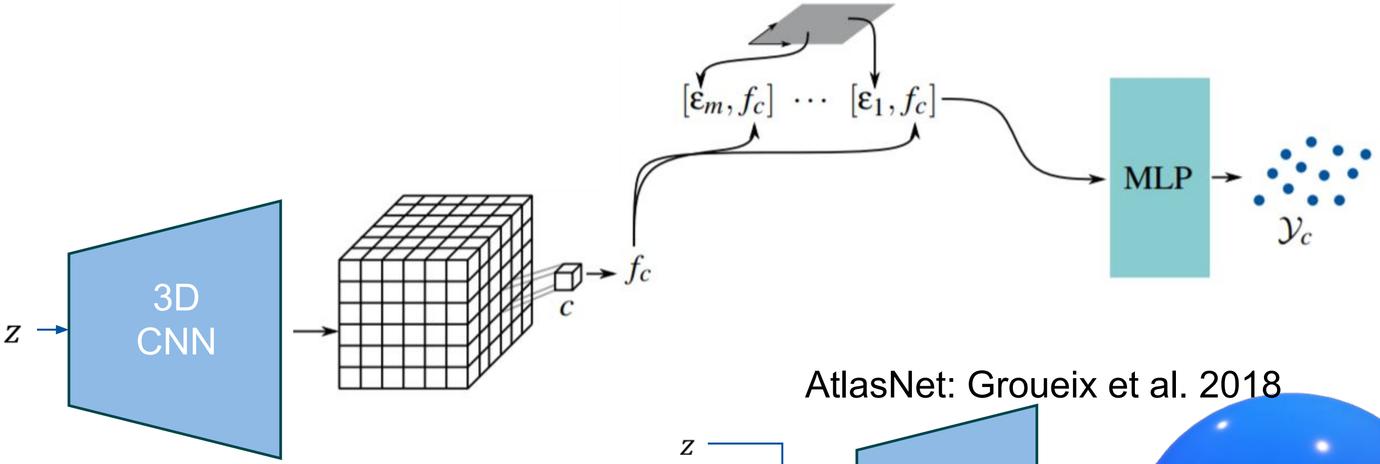


AtlasNet: Groueix et al. 2018

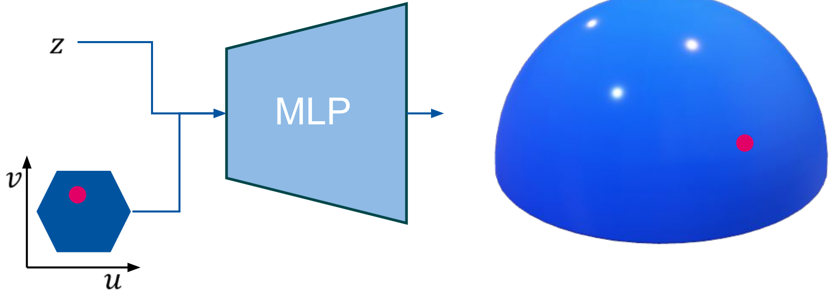


Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

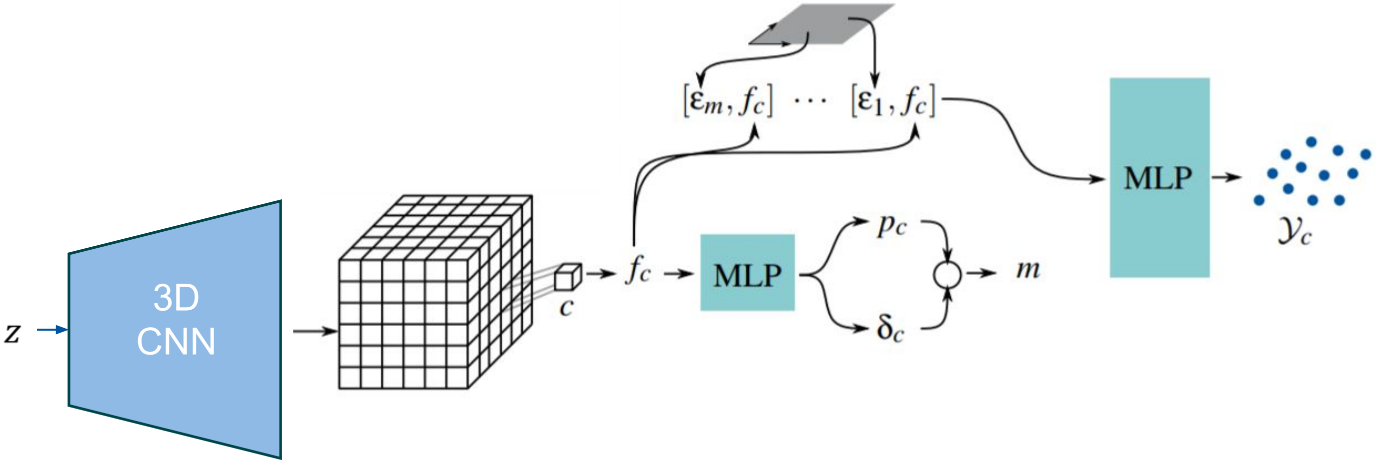
Learning Localized Maps



AtlasNet: Groueix et al. 2018

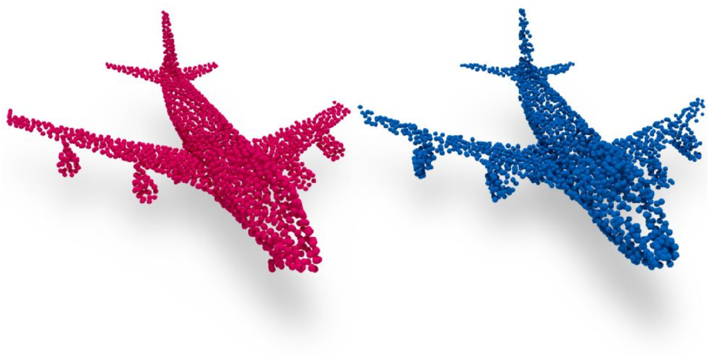


Learning Localized Maps



Visual Computing Institute | Isaak Lim
Computer Graphics and Multimedia

Variable number of output points



input: 2500

output:
2500

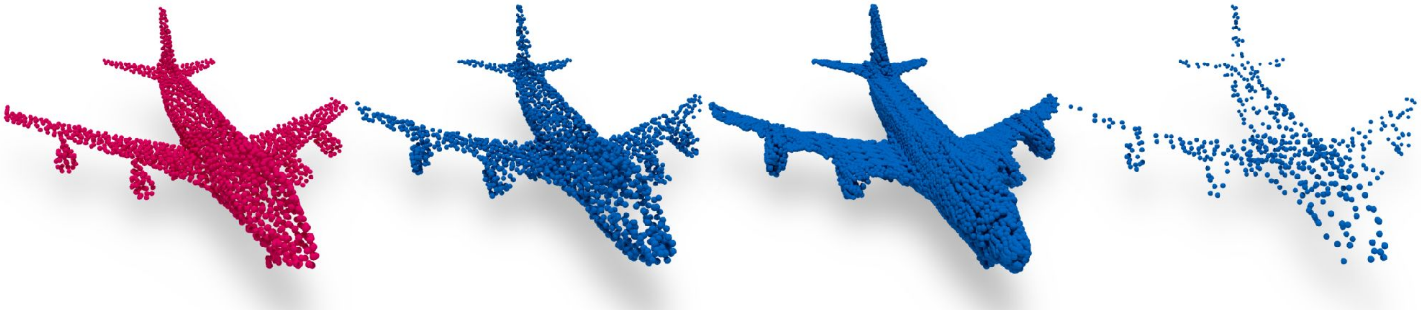
Variable number of output points

input: 2500

output: 2500

output: 15000

Variable number of output points



input: 2500 | output: 2500 | output: 15000 | output: 500

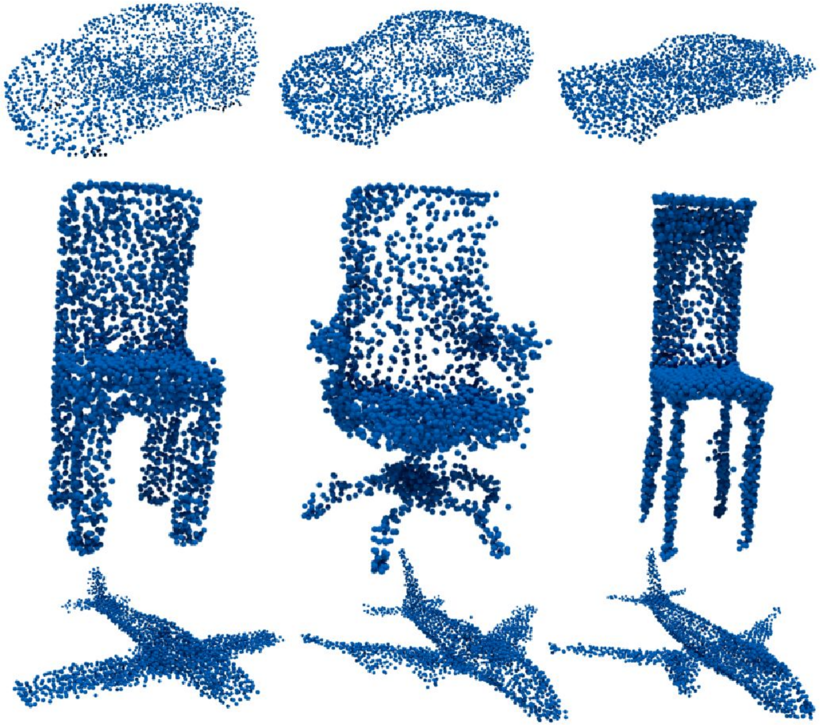
Variable number of output & input points

input: 50 output: 16000 ground truth: 16000

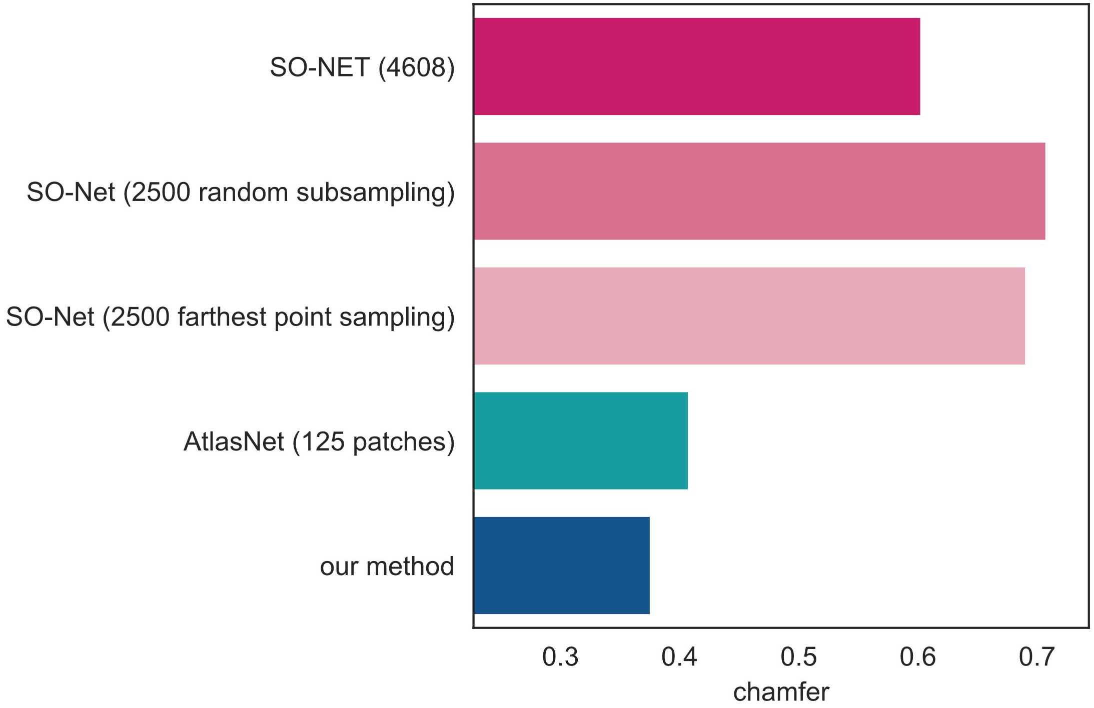
Variable Encoders

Input output ground truth

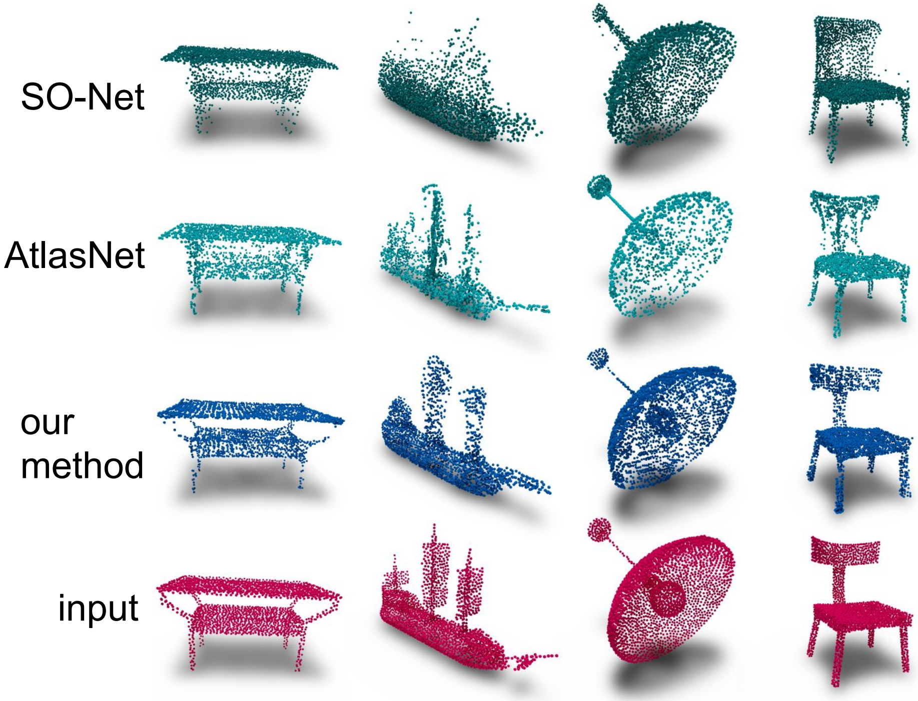
Variable generation of z



Comparison: our dataset



Comparison: our dataset



Key Questions

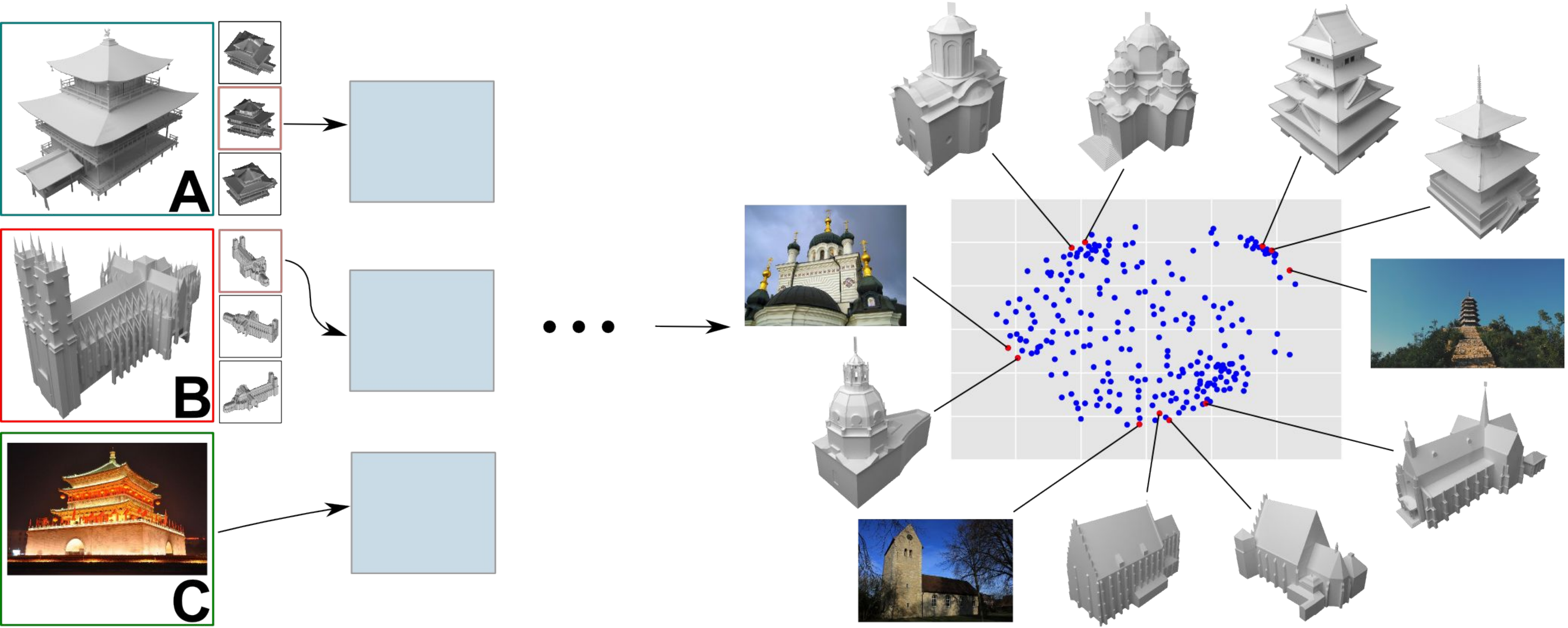
- How can we compute encodings of geometric data?
- **How can we design the embedding space?**

Our contribution

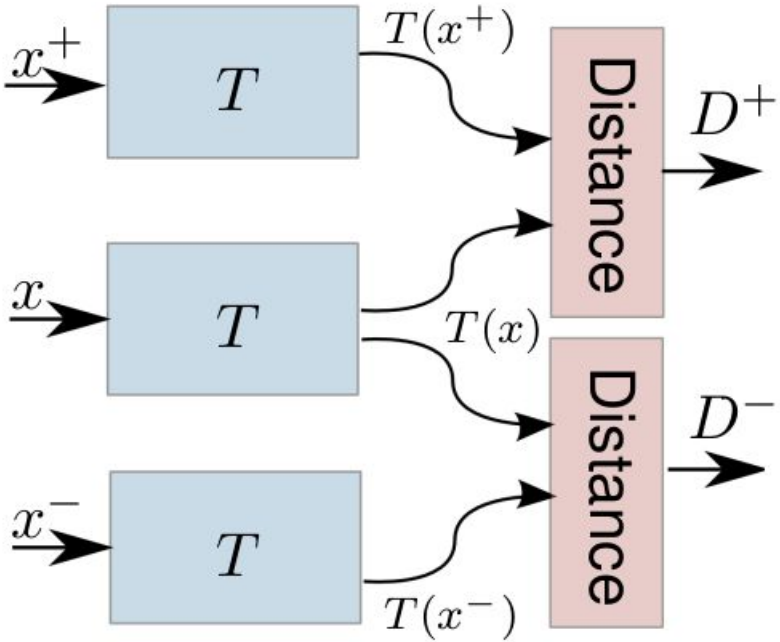
Identifying Style of 3D Shapes using Deep Metric Learning

Metric Learning

Metric Learning

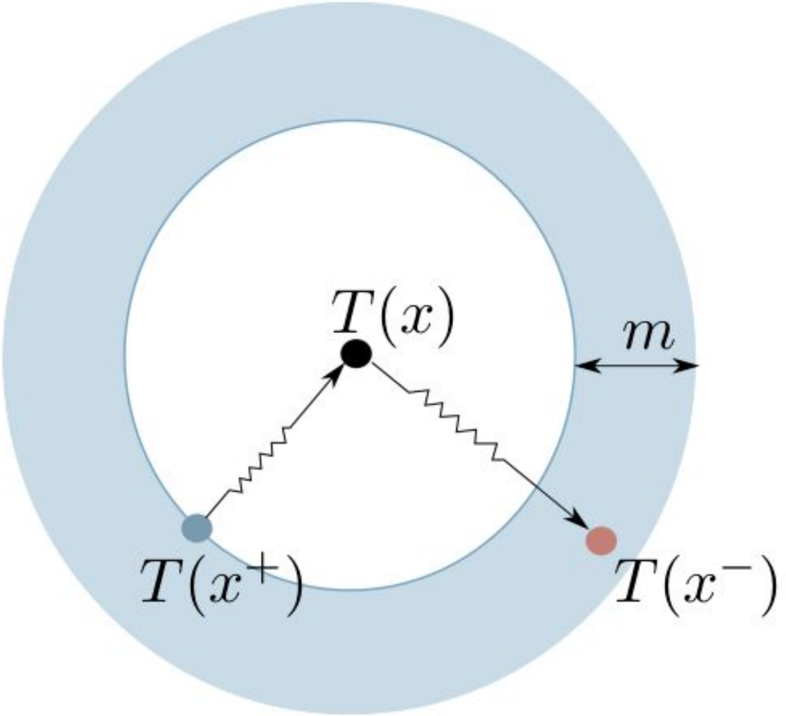


Metric Learning



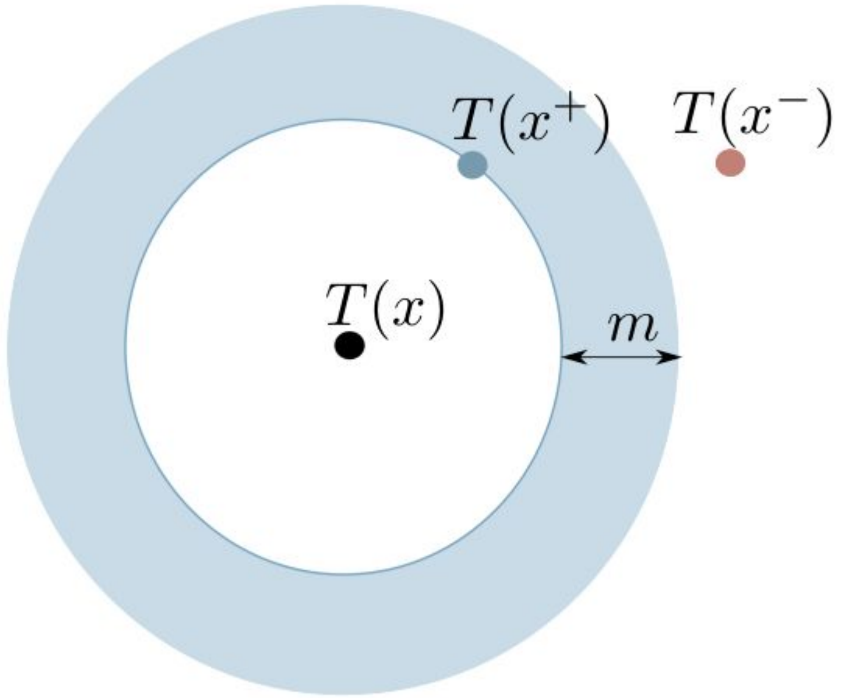
Metric Learning

$$L(x, x^+, x^-) = \max(0, \boxed{m + D^+(x, x^+) - D^-(x, x^-)})$$

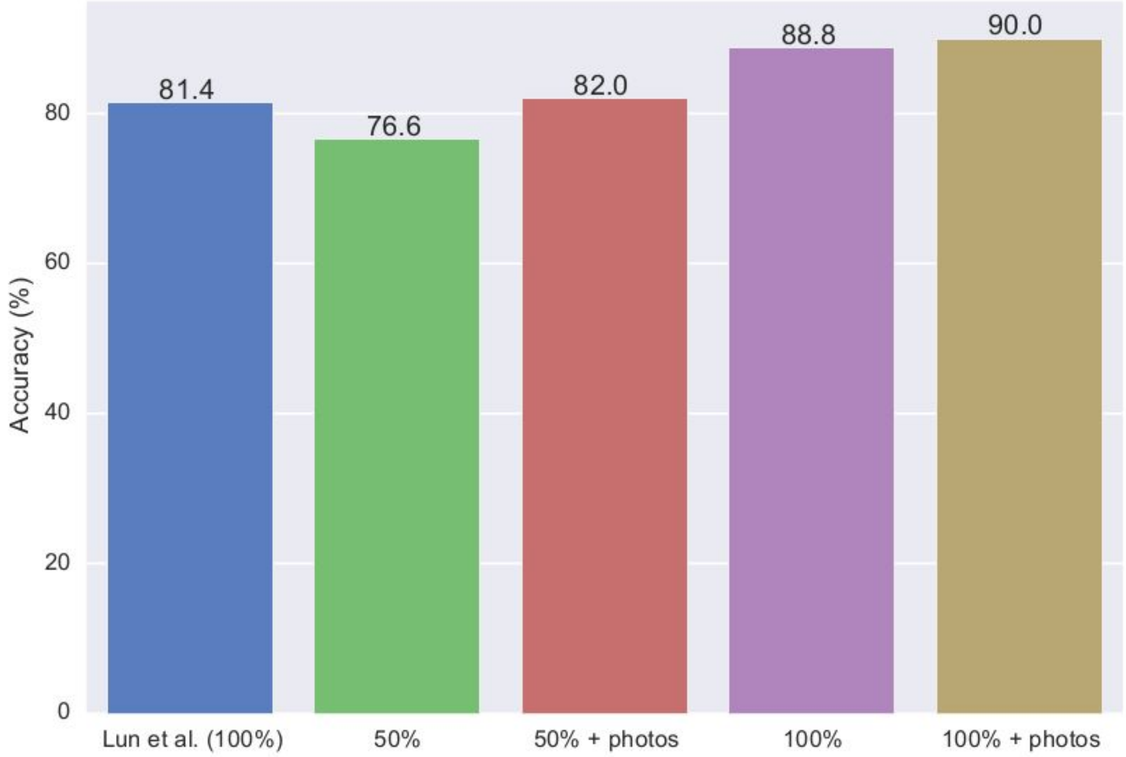


Metric Learning

$$L(x, x^+, x^-) =$$
$$\max(0, m + D^+(x, x^+) - D^-(x, x^-))$$



Metric Learning



Conclusion

- Encodings of geometric data
- Designing the embedding space
- Meaningful combinations of representations for different tasks